Gibbsian properties

Ermolaev, Victor Nikolaevich

1

Introduction

One may think of the starting point of the theory of Gibbs measures as an (successful) attempt to answer the question about physical reality: “How can a ferromagnet or a gas or a liquid in thermal equilibrium be described in mathematical terms?”. The main feature to be captured is the existence of different *phases* of a system, for instance positive and negative magnetizations for magnets, existence as a liquid, steam, or ice for water. However, this question could be extended to a whole class of physical systems by identifying the common features of a ferromagnet, a gas, and a liquid. The most straightforward and promising characteristic is that all of them contain an enormously large number of microscopic components taking values in the same state-space. Therefore the question may be transformed into another: “How can a system consisting of large number of identical interacting components in equilibrium be described in mathematical terms?”. The first thing we do, to simplify, is replacing physical space by a suitable graph. As an example one may think of a lattice , which is actually physically realistic for crystals, but it can also be used it as a simplified model for continuous space, or a tree (e.g. a Bethe lattice) as an even cruder model, although it is often appropriate for models met in biological applications. Despite the relative simplification, the microscopic structure of such a system is still extremely complex, while its macroscopic characteristics are often not very complicated (temperature or pressure for a gas, magnetization for a ferromagnet, etc.) The idea of Maxwell, Boltzmann, and Gibbs rephrased in modern mathematical language was as follows: the microscopic complexity can be overcome by a statistical approach; the macroscopic determinism may be regarded as a consequence of an application of a suitable law of large numbers. So to say, no state of a system should be described by a fixed configuration of the system's components. This description should rather be given in terms of a family of random variables $\{\sigma_i\}$. These random variables are associated with the *sites* i of the graph chosen as a mathematical abstraction of a physical space. Consequently, the joint distribution μ of $\{\sigma_i\}$ determines the system's state. The finite-volume prescription is given by Gibbs' canonical ensemble prescription

$$\mu(d\sigma) = \frac{1}{Z} e^{-\beta H(\sigma)} d\sigma, \tag{1.1}$$

where the function $H(\sigma)$ refers to the energy of any configuration σ — the *Hamiltonian* — and β is an inverse temperature. $H(\sigma)$ is formed by the interactions of all microscopic components and (perhaps) an external force h. In the case of ferromagnet, h has a meaning of an external magnetic field. The finite-volume Gibbs measure $\mu(d\sigma)$ is always unique, therefore there exists only one phase for any finite-volume system.

The next step in achieving the goal of an adequate description of existence of several phases is to identify large-volume (physically) or infinite-volume (mathematically) behaviour. It is easy to see that the formula (1.1) is ill-defined for any infinite configuration. This is due to the fact that the energy of an event in infinite volume is also infinite. There are two approaches developed to overcome this difficulty. The first one is expressing energy of a configuration as a function of its size and straightforward taking limits as the volume grows. Under some suitable conditions such limit(s) exist(s). Existence of several limiting measures will correspond to a *phase transition*. The second approach is based on the very fruitful idea that instead of performing this limiting procedure, one may study directly the set of all possible limiting objects and their *conditional restrictions* on finite sub-volumes fixing the configuration outside of them. These restrictions are given by the finite-volume prescription now depending on the changing inside and the fixed outside of particular volumes in some suitable way (this issue will be discussed shortly). Each extremal limiting object corresponds to a phase of a system. The family of the finite-volume conditional distributions may be compatible with more than one limiting measure, then a phase transition occurs, this phenomenon has been widely studied in many articles and books.

The former discussion can be summarized in two ways: physically, *a Gibbs measure is a mathematical idealization of an equilibrium state of a physical system which consists of a very large number of interacting components*, mathematically *a Gibbs measure is the distribution of a stochastic process which, instead of being indexed by time, is parametrized by the spatial coordinates of the system and has the special feature of admitting prescribed versions of the conditional distributions outside finite regions*, Georgii [26].

To be more specific we consider a model living on a lattice. Let $\mathbb{L}$ be the whole lattice and Λ its finite-volume subset. Let σ_i be a random variable (a *spin*) sitting at site i of $\mathbb{L}$ and taking values from S, then the joint distribution of all spins σ_Λ in finite volume Λ takes values in $S^\Lambda =: \Omega_\Lambda$, the probability space of the whole system is therefore $S^{\mathbb{L}} = \Omega$. Within this mathematical framework the question whether a family of conditional probabilities $\{\gamma_\Lambda, \Lambda \Subset \mathbb{L}\}$, a *specification*, is Gibbs arises naturally. A beautiful theorem was proven by Sullivan [54] and later extended by Kozlov [35], stating that uniform non-nullness and quasilocality properties are necessary and sufficient conditions for a specification to be Gibbs. Uniform non-nullness refers to the fact that all conditional probabilities in the family of any event are separated from zero by a positive constant. Quasilocality is a property that guarantees that the energy function is well-defined, differently this property may be expressed as follows:

$$\sup_{\omega,\eta,\tilde{\eta}\in\Omega} |\gamma_\Lambda(\sigma_\Lambda|\omega_{\Lambda_n}\eta) - \gamma_\Lambda(\sigma_\Lambda|\omega_{\Lambda_n}\tilde{\eta})| \xrightarrow{n\to\infty} 0, \tag{1.2}$$

for any ω. In mathematical terms, (1.2) corresponds to the fact that all intermediate configurations ω shield σ_Λ from the influence of the outside of ω; in physical terms, this statement has a more intuitive form: quasilocality ensures that any local experiment is controllable. Imagine a well-tuned guitar and an astronaut playing it in a space ship. No matter if a space ship is being prepared to start its journey from Earth to outer space or if a space ship has already started and reached the vacuum of the cosmos, the guitar will sound the same. The exterior of a space ship plays no role. The interior of a space ship filled with air (or close to air mixture of gases) creates a barrier preventing the influence of either vacuum or air outside of a space ship. Hence, playing a guitar is quasilocal experiment.

If the spin space is discrete, as it will mostly be in this thesis, quasilocality equals continuity (in the product topology).

While non-nullness of a specification is not a property to be ignored, it is not usually the main problem. The lack of quasilocality leads to more serious consequences, namely in this case the Boltzmann-Gibbs prescription is not any more well-defined. In heuristic terms and in the light of (1.2), non-quasilocality means that there is some mechanism transmitting the influence of "far-away" regions to any finite sub-volume and remaining active even in the case of no fluctuations in the spins in-between. For a measure to be non-Gibbs it is enough that there exists only one such configuration of the in-between spins. This very idea is the main tool of the Gibbs-non-Gibbs investigation.

Now let us explain why the question whether or not a given measure was a Gibbs measure first arose. This happened in the theory of equilibrium statistical mechanics, in the study of phase transitions. To study the behaviour of a system close to critical points or to prove the absence of phase transitions for a system, a *renormalization map* $\mathcal{R}$ was applied to the Hamiltonian of system's Gibbs measure μ, [27, 28]. Although such maps were studied before, these papers for the first time put the renormalization-group ideas and the Gibbs formalism together. The absence of phase transitions at high temperature and/or low density was known before, and the analysis in this region served as an illustration. The map $\mathcal{R}$ is expected to define a new Hamiltonian $H' := \mathcal{R}(H)$ and moreover induces a map K mapping the original measure μ to μ'. Having K defined, the existence of H' depends on whether the image measure μ' is Gibbs or not [11, 33]. If the image measure μ' is Gibbs, then the associated Hamiltonian H' is well-defined, otherwise — not. It was found that the application of many maps may result in such "pathologies" [12, 19, 41]. When such a phenomenon appears it means that finite-volume conditional probabilities of the image system will acquire long-range dependencies (therefore lose its quasilocality property), at least for some non-removable configurations.

After having discovered this, another question attracted a lot of interest and attention: "How stable is the Gibbs property of a measure describing a system in equilibrium, when the last is subjected to a time evolution?". There is a long expanding list of references on this: starting from spin contraction [9, 44], general single-site transformations [38], continuing with stochastic spin evolutions for bounded and unbounded spins in [7, 12, 14, 36, 37, 39, 48] for different kinds of underlying graphs of a system, and going further. A good example of losing Gibbsianness under time evolution is given in [12]. Before explaining it, we need to discuss informally the celebrated Ising model, a paradigmatic statistical mechanical model, designed to study magnetic properties with spins taking values in $\{-1,+1\}$. We choose as time evolution the *spin-flip* (or *Glauber*) dynamics: each spin is flipped according to independent Poissonian clocks attached to a site. Rephrasing the proof in [12, see Theorem 6.3, fact 2], we start with the Ising model on lattice. Such a model at low temperature will typically be found in one of two phases (low-temperature extremal Gibbs measures) , which look similar to a *ground* state – state having the biggest probability to be observed – almost everywhere pluses or almost everywhere minuses. We start with a low-temperature Gibbs measure. We choose a site 0, which will be called the *origin*, let the time flow and flip a spin when an associated to it clock rings. After some time a configuration $\omega_{\Lambda\setminus 0}\sigma_{\Gamma\setminus\Lambda}$ will be observed, where Λ is a finite subset of Γ with the property that Γ is a lot larger than Λ. We suppose that the configuration $\omega_{\Lambda\setminus 0}$ is *atypical* (or, very improbable) and $\sigma_{\Gamma\setminus\Lambda}$ is typical for any of the possible starting phases. Though the configuration in $\Gamma\setminus\Lambda$ may be not typical for the phase we started with, the cost of creating such a configuration is proportional to the length of contour separating Γ from the rest of the lattice. On the other hand, the cost of inserting any atypical configurations in $\Lambda\setminus 0$ is proportional to the volume of Λ. Suppose $\Lambda\setminus 0$ is not too small, that it could be considered as a barrier shielding the origin from the influence of $\sigma_{\Gamma\setminus\Lambda}$. Nonetheless, the effectiveness of such a barrier is poor because the values of the spin at the origin at time 0 and later time are positively correlated. Thus, there exists a time window (possibly infinite) with values such that they are large enough for the dynamics to create an atypical configuration around the origin and not too big to keep the correlation between the values of spin at the origin at different times. Within this time interval $\sigma_{\Gamma\setminus\Lambda}$ determines the value of σ_0 and the quasilocality is lost.

The general picture is that for very general dynamics and very general initial measures the time-evolved measures are again Gibbsian, for a sufficiently small time-interval [10, 12, 37, 41, 43]. Long times however, even for simple dynamics offer the possibility for the emergence of non-Gibbsian measures. The discontinuities in the conditional probabilities which are responsible for the Gibbs-non-Gibbs transitions are produced by hidden phase transitions which pop up as a result of the conditioning procedure. Depending on the specific nature of the system there may be many mechanisms of such singularities [12, 36].

The present thesis attempts to investigate when the Gibbs property of an initial measure is preserved or lost if the initial Gibbs model is subjected to some transformation. This question is addressed for two classes of statistical mechanics models. The first class is mean-field models, where the underlying structure is a complete graph, so each component influences on all the others. The simplifications made for mean-field models allow to develop tractable results. The importance of these models is due to their ability to mimic the behaviour of large-dimensional lattices [37, 52], as well as their tractability. Mean-field models are objects of interest on their own in computer science as they catch well the behaviour of computer networks [2, 40], for example. We warn the reader that in the mean-field case, the notion of Gibbsianness will be slightly changed. This will be discussed later, and in full detail in Chapter 4. The second class concerns models on trees, where components interact in a local fashion. Tree models form a first step way from a mean-field setup towards a proper lattice model. The main achievement of this thesis is a rather complete description of the Gibbs-non-Gibbs regions in the space of temperature and time followed by explicitly presented equations governing the dynamics in the mean-field setup in vanishing external field.

In the end of the present thesis we will address a problem of classification coming from Information Theory and show how to approach it with the Gibbsian formalism.

1.1 Strategy

Following the route suggested in [12], the study of a failure of the Gibbs property under stochastic time evolution is connected with understanding of a “constrained” or “two-layers” model. This model reflects the influence of a dynamics. The effect of evolution on the initial Gibbs measure results in some transformations of the energy function of its finite-volume restrictions. The lack of phase transitions in “constrained” models immediately guarantees the Gibbs nature of the evolved system. On the other hand the existence of phase transitions may or may not imply the non-Gibbsian nature depending on the setup, often an extra step to show the lack of the quasilocality property is needed.

For tree models, as for more general lattice models, quasilocality means that the influence of "far-away" regions on any finite regions is effectively stopped by the spins in-between, in other words, the interaction between spins has a local nature. Thus, the conditioning on "middle-distanced" spins is essentially the same as the conditioning on them and any "far-away" configuration. This implies that the finite-volume probabilities of the transformed measure conditioned on configurations different only on "far-away" regions should be almost equal. If not — the "middle-distanced" configuration is called *bad* and quasilocality is lost. For discrete spins, which are of interest in the present thesis, the property of quasilocality is equivalent to the continuity property of finite-volume conditional probabilities as a function of the conditioning (in the product topology).

The complete graph, which is used in mean-field models, discards the notion of "far-away" regions, because all spins communicate with each other equally, so the concept of quasilocality is meaningless in this setup. The presence of a phase transition in the transformed system still plays a role, however. Due to the graph structure, any configuration and all its permutations may be identified with a *real number*, hence conditioning on a configuration translates in conditioning on this real number. The main characteristic of conditioning becomes a *magnetization* (or, more generally, some form of an *empirical* average) of a system. A mean-field system is called non-Gibbsian if the single-site conditional probabilities depend in a discontinuous way on the magnetization of the conditioned spins. The last statement suggests an existence of a phase transition is a sufficient condition for non-Gibbsianness.

Hence, to deduce a lack of Gibbsianness for the transformed system one is left with the problem of investigating continuity properties of conditional probabilities for the "constrained" model.

1.2 Overview of Thesis

In this section we provide an informal guide to what is contained in the main body of the thesis and outline the main achievements of the present work.

In Chapter 2 we review the formal side of the theory of Gibbs measures, making rigorous statements and notions used in the introduction. Chapter 2 provides a common description for many models of statistical mechanics. A broad description of dynamics is given, and the question of how applying a dynamics can cause the loss of the Gibbsian property for the evolved (according to this dynamics) measure starting from an initial Gibbsian measure is addressed. A general strategy of proving non-Gibbsianness is described in detail.

In Chapter 3 we look at tree models. In particular, we consider Cayley trees for simplicity. As explained in Chapter 3, an Ising model on a Cayley tree exhibits a phase transition, therefore several (among which there are three extremal homogeneous) Gibbs measures are admitted. We discuss when a time-evolved initial Gibbs measure loses its quasilocal property. All three measures will be shown losing the quasilocality property under time evolution after some time. Moreover, we keep track of *how many* configurations transmit an influence of "far-away" regions. Surprisingly, for one of the initial measures after some time *all* configurations become bad. This fact holds independently of the preference of the system to be in a certain state, but the preference value has to be smaller than a certain real number. The preference of the system to a state is expressed by an external magnetic field aligning the spins. At the end of Chapter 3 we point out possibilities to extend the results for Cayley trees to a more general class and explain why this has to be possible by identifying common tokens of Cayley and general trees. Chapter 3 has appeared as [13]. Chapter 3 positively answers on the question whether non-Gibbsianness becomes worse as time progresses in the Cayley tree setup.

Chapter 4 is devoted to the description of the time evolution of mean-field systems in thermal equilibrium subjected to arbitrary-temperature Glauber dynamics. The results of Chapter 4 have appeared as [15] and extend the work of Külske and Le Ny [36] contributing an analysis of low-temperature dynamics. The low-temperature dynamics corresponds to a *dependent* spin-flip dynamics. A detailed description of the decomposition of the time-temperature space into Gibbs and non-Gibbs regions is given. Each couple of coordinates in the time-temperature plane depending on its location will correspond to a Gibbs/non-Gibbs regime of the evolved system. In Section 4.2 of Chapter 4 we formulate the main result which covers the up-to-now accumulated knowledge of arbitrary-temperature dynamics for mean-field models.

Chapter 5 demonstrates an interplay between the Gibbsian formalism and applied problems. In particular, we show how the theory of Gibbs measures can be involved in solving classification problems. Such problems are of great importance, for example, in biological science and Information Theory. While solving classification problems or problems of denoising, one is usually given random samples and aims to (re-)construct a probability distribution which generated these samples. We build our analysis based on the simple observation that the nature of the distribution of interest is intrinsically Gibbsian. This notice allows us to develop an approach related to the potential of a Gibbs measure. We show that this way of treating problems can successfully solve problems of classification.

* * *